

# Selection of Immunogenic Peptides for Antisera Production

In order to produce antisera reactive with proteins, a peptide is selected from a translated cDNA or protein sequence and is then synthesized, purified (at least partially), and conjugated to a carrier protein. Strategies for selecting immunogenic peptides and carrier proteins as well as various means of coupling peptides to carriers are described. The actual coupling protocols, a typical immunization protocol, and a procedure for preparing a peptide affinity column for antibody purification are included in UNIT 9.4. A flowchart for the preparation of antipeptide antibodies (antibodies reactive with a synthetic peptide) is shown in Figure 9.3.1.

## SELECTION OF AN IMMUNOGENIC PEPTIDE

To prepare antibodies against synthetic peptides, a peptide must first be selected. This is the most critical step in obtaining an antibody

that reacts with the native antigen. In practice, a 10- to 15-residue peptide sequence inferred from a cDNA sequence or from an N-terminal amino acid sequence is selected. If possible, sequences should be avoided that are likely to be identical or highly homologous to those in the animal to be immunized (usually rabbits). After synthesis and purification, the peptide is cross-linked to a carrier protein such as keyhole limpet hemocyanin (KLH).

If a peptide sequence to be utilized for antisera production is not from the terminal regions of the protein, selection is based on predicting antigenic sites. It is presumed that the sites accessible to reactivity with antipeptide antibodies are exposed on the surface of the protein; these sites are likely to be more common in flexible regions of the protein (Westhof et al., 1984), and are more likely to be found on reverse turns or loop structures (Dyson et al., 1985). Computer algorithms

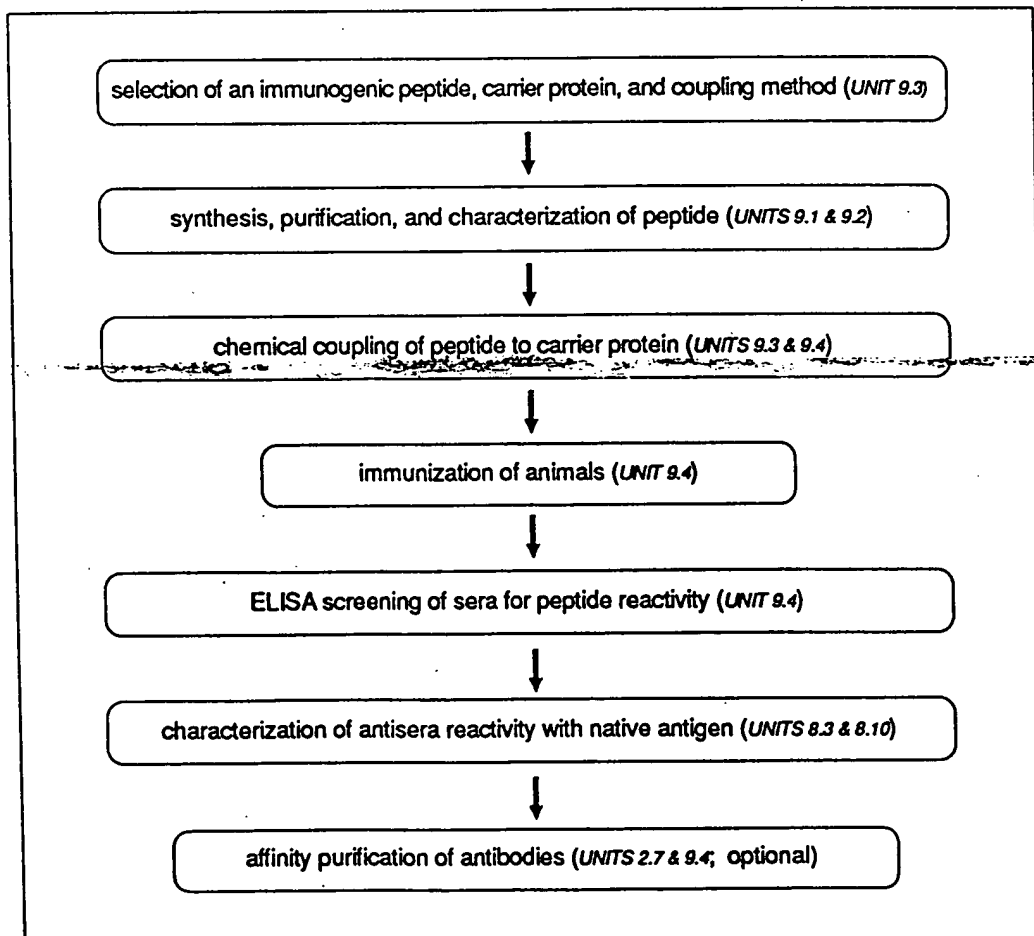


Figure 9.3.1 Flow chart for preparation and analysis of antipeptide antibodies.

estimate hydrophilicity (Kyte and Doolittle, 1982; Hopp and Woods, 1983), flexibility (Karplus and Schultz, 1985), and secondary structure of the protein along its polypeptide chain; they can help identify surfaces on the protein that have appropriate characteristics for antigenic sites. Antisera produced against peptides derived from such regions are more likely to react with the native protein than peptides derived from regions without these characteristics.

### Selection of a C-Terminal Peptide

Because the C termini of proteins are often more mobile than the rest of the molecule and are frequently exposed on the protein surface, this region is usually accessible to antibody-combining sites. This is particularly true for detergent-solubilized transmembrane proteins with C-terminal cytoplasmic tails—i.e., MHC class I and class II molecules, and T cell receptor molecules. This type of peptide can be coupled to the carrier in a straightforward manner using *m*-maleimidobenzoyl-*N*-hydroxy-succinimide ester (MBS) via a Cys residue that has been added to the N terminus of the selected peptide. By coupling the peptide via its N-terminal end to the carrier protein, the peptide will be exposed in a fashion similar to that found in the native antigen. Obviously, if any other Cys residues are present, they will also couple; therefore, an alternate coupling procedure may be more appropriate (see coupling methods).

Unfortunately there will be instances when the C-terminal sequence will not be the preferable immunogen—e.g., if its sequence is part of a transmembrane region and is thus too hydrophobic. This can be determined by examining a hydropathicity plot of the sequence (Kyte and Doolittle, 1982).

### Selection of an N-Terminal Peptide

As with the C terminus, the N terminus is often exposed in the native protein; therefore, peptides selected from this region are often useful for making antibodies reactive with the intact protein. In this case, if MBS coupling is used, the Cys residue should be present on the C terminus. By coupling the peptide via the C-terminal end to the carrier protein, it will be oriented as it is in the native protein. Unfortunately, the N terminus of a protein may have a post-translational modification, such as acetylation or fatty acid acylation. Moreover, if the protein sequences are derived from cDNA sequences, the leader sequence must not be confused with the authentic N terminus. It is usu-

ally possible to locate potential leader sequences using an algorithm derived by von Heijne (1986). If it is known that the authentic N terminus is acetylated, the peptide can then be acetylated during synthesis to reproduce the structure in the native protein.

### Selection of an Internal Peptide Sequence

Selection of a peptide from an internal part of the protein sequence can be aided by the use of algorithms to predict those regions most likely to be exposed on the surface of the protein. The only information needed is the primary amino acid sequence. Two of the methods discussed below are based on the fact that antibody-reactive sites are usually located in externally exposed, hydrophilic regions of proteins. A third method relies on secondary structure predictions.

Using the first algorithmic method, a hydrophilicity value is assigned for each overlapping six-amino-acid segment of the protein sequence based on the average of the hydrophilicity values (Table 9.3.1) of the amino acids in that segment. The highest point of average local hydrophilicity is usually located in or near an antigenic determinant. A computer program written in Basic is available for analysis (Hopp and Woods, 1983).

An alternative algorithm evaluates the hydrophobic and hydrophilic tendencies of a polypeptide chain based on water vapor free-energy transfers and the interior versus exterior distributions of amino acid side chains. Values for each amino acid are listed in the second column of Table 9.3.1. A computer program for calculating this hydropathicity profile has been written by Kyte and Doolittle (1982). This profile is useful for determining exterior and interior regions of a native protein, as well as for locating signal sequences and transmembrane sequences.

A third algorithm is an empirical method that relies on a library of known structures to determine the frequency with which each amino acid occurs in the various conformational states (i.e.,  $\alpha$ -helix,  $\beta$ -sheet,  $\beta$ -turn, or all other structural forms; Chou and Fasman, 1974). Using these frequencies, predictions can be made about secondary structure for a given sequence. For making anti-peptide sera, regions that are predicted to form turns or loops, or extended sequences (20 to 25 residues) that have a very high probability for formation of an  $\alpha$ -helix, are useful. A computer program for performing this analysis can be found in

Corrigan and Huang (1962).

Although no data exists to prove this point, it would seem wise to avoid choosing peptides containing predicted polysaccharide attachment sites, most notably the sequence Asn-X-Ser or Asn-X-Thr, which predict the presence of an Asn-linked polysaccharide. It is likely that the presence of polysaccharide moieties at such sites in the native protein would interfere with antibody accessibility.

The above-mentioned computer programs, plus programs to predict flexibility, location of transmembrane regions, Asn-linked glycosylation sites, and sites of signal sequence cleavage, are all contained in a package called PC Gene produced by IntelliGenetics (APPENDIX 5).

### Selection of the Length of the Peptide

Generally, peptides with a length of 10 to 15 residues are used to make anti-peptide sera that react with the native protein. Peptides with as few as 6 or as many as 35 amino acids have worked successfully; however, both extremes have disadvantages. Small peptides are more soluble and can produce very specific antisera, but the antibodies elicited by them are not as likely to react with the parent protein. Large peptides tend to be less soluble, more difficult to prepare synthetically, and are more likely to

assume structures unrelated to the native protein. Part of the decision about peptide size will be determined by the individual peptide sequence, as some residues will adversely affect solubility, produce synthesis problems, or interfere with coupling.

Thus, in summary, a reasonable order of suggestions for choosing peptide sequences for making anti-peptide sera would be:

1. If possible, use more than one peptide.
2. Use the C-terminal sequence (7 to 15 residues) if it is hydrophilic and if a suitable coupling group is available or can be added.
3. Use the N-terminal sequence (7 to 15 residues) if it is hydrophilic and if a suitable coupling group is available or can be added.
4. Use internal hydrophilic regions, perhaps using longer peptides (15 to 20 residues).

### Modification of the Chosen Peptide

Other features of the peptide must be considered in order for it to mimic the native antigen as closely as possible. If the desired peptide sequence comes from an internal portion of the native protein, then the free N-terminal amino and C-terminal carboxyl groups (which are normally peptide-bonded to adjacent amino acid residues within the native protein) can be modified to more closely

Table 9.3.1 Hydrophobic/Hydrophilic Index of Amino Acids

Amino acid	Hydrophilicity value <sup>a</sup>	Hydropathy index <sup>b</sup>
Arginine (R)	3.0	-4.5
Aspartic acid (D)	3.0	-3.5
Glutamic acid (E)	3.0	-3.5
Lysine (K)	3.0	-3.9
Serine (S)	0.3	-0.8
Asparagine (N)	0.2	-3.5
Glutamine (Q)	0.2	-3.5
Glycine (G)	0.0	-0.4
Proline (P)	0.0	-1.6
Threonine (T)	-0.4	-0.7
Alanine (A)	-0.5	1.8
Histidine (H)	-0.5	-3.2
Cysteine (C)	-1.0	2.5
Methionine (M)	-1.3	1.9
Valine (V)	-1.5	4.2
Isoleucine (I)	-1.8	4.5
Leucine (L)	-1.8	3.8
Tyrosine (Y)	-2.3	-1.3
Phenylalanine (F)	-2.5	2.8
Tryptophan (W)	-3.4	-0.9

<sup>a</sup>Hopp and Woods (1981).

<sup>b</sup>Kyte and Doolittle (1982).

Tabl 9.3.2 Principle Carriers Used for Coupling Peptides<sup>a</sup>

Carrier <sup>b</sup>	$M_r$ (kDa)	Number of groups/molecule			
		$\epsilon$ -NH <sub>2</sub>	-SH	Phenol	Imidazole
BSA	67	59 <sup>c</sup>	1	19	17
Ovalbumin	43	20	4	10	7
Myoglobin	17	19	0	3	12
Tetanus toxoid	150	106	10	81	14
KLH	>2000	6.9 <sup>d</sup>	1.7 <sup>d</sup>	7.0 <sup>d</sup>	8.7 <sup>d</sup>

<sup>a</sup>Adapted from Van Regenmortel et al. (1988).

<sup>b</sup>Abbreviations; BSA, bovine serum albumin; KLH, keyhole limpet hemocyanin.

<sup>c</sup>Only 30 to 35 of the 59 Lys residues of BSA are accessible.

<sup>d</sup>For KLH, the amino acid groups are expressed in grams of amino acid containing this functional group per 100 g.

mimic their native structure. The N-terminal amino group can be modified by acetylation of the peptide  $\alpha$ -amino group during synthesis, and the C-terminal carboxyl group can be modified with a C-terminal amide during peptide synthesis. It is not certain that acetylation of the N terminus or formation of the C-terminal amide for peptides derived from internal sequences will really improve the chances of producing antisera reactive with the native protein. However, it has been demonstrated that these modifications will stabilize an  $\alpha$ -helical conformation and may increase the solubility of the peptide.

### SELECTION OF A CARRIER PROTEIN

A carrier protein should be a good immunogen and have a sufficient number of amino acid residues with reactive side-chains (see Table 9.3.2) for coupling to the synthetic peptide. KLH is commonly used because of its proven efficacy. However, because of its large size, KLH is more likely to precipitate during cross-linking, making the complex difficult to handle. Other proteins that have been used as carrier molecules include thyroglobulin, bovine serum albumin (BSA), and tetanus toxoid. BSA has the disadvantage that anti-carrier protein antibodies present in the anti-peptide serum will be a problem if the antiserum is used in the presence of fetal calf serum. Table 9.3.2 lists the most common protein carriers and their relevant properties.

### SELECTION OF A COUPLING METHOD

In addition to the choice of peptide, a method for coupling the peptide to a protein carrier must also be selected. Most coupling methods rely on the presence of free amino

(Lys), sulfhydryl (Cys), phenolic (Tyr), or carboxylic (Asp or Glu) groups. The chosen coupling method should link the peptide to the carrier via either the C- or N-terminal residue. Peptides corresponding to the amino terminus of proteins should be coupled through their carboxyl-terminal amino acid residue, whereas peptides corresponding to the carboxyl terminus of proteins should be coupled through their amino-terminal amino acid residue.

One coupling procedure that has proved to be particularly effective employs MBS as the coupling reagent. This procedure requires a free sulfhydryl group on the synthetic peptide and free amino groups on the carrier protein. Therefore, in order to use this method, it is usually necessary to add a Cys residue (during peptide synthesis) to the C or N terminus of the peptide. This will provide the sulfhydryl group for coupling to the carrier protein. Coupling a peptide derived from an N-terminal sequence to a carrier is accomplished with MBS via a Cys residue added to the C terminus of the peptide. This cross-links the peptide derived from the N-terminal sequence to the carrier molecule so that the N terminus is exposed as it would be in the native antigen. For a peptide derived from a C-terminal sequence, the Cys is placed on the N-terminus of the peptide for the same rationale. If an internal Cys residue is present in the chosen peptide, it can be used for MBS coupling, especially if the cysteine is part of a disulfide linkage in the native protein.

If coupling with MBS is not desirable for some reason (such as the presence of non-terminal Cys residues that are not disulfide-linked in the native protein), coupling can then be accomplished through a C- or N-terminal Tyr using *bis*-diazotized benzidine (BDB), through the C or N terminus with 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDCI), or

Tabl 9.3.3 Principal Reagents Used for Peptide Protein Conjugation<sup>a</sup>

Coupling agent	Modified amino acid	
	Primary reaction	Secondary reaction
Glutaraldehyde	$\epsilon$ -NH <sub>2</sub> Lys, $\alpha$ -NH <sub>2</sub> , SH-Cys	Tyr, His
Bis-imido esters	$\alpha$ -NH <sub>2</sub> , $\epsilon$ -NH <sub>2</sub> Lys	Negligible
BDB	Tyr, SH-Cys, His, $\epsilon$ -NH <sub>2</sub> Lys	Trp, Arg
Carbodiimides (EDCI)	$\alpha$ -NH <sub>2</sub> , $\epsilon$ -NH <sub>2</sub> Lys, $\alpha$ -COOH, Glu, Asp	Tyr, Cys
MBS	Cys-SH	Not observed

<sup>a</sup>Adapted from Van Regenmortel et al. (1988). Abbreviations: BDB, *bis*-diazotized benzidine; EDCI, 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide; MBS, *m*-maleimidobenzoyl-*N*-hydroxysuccinimide.

through the N-terminal  $\alpha$ -amino group with glutaraldehyde. BDB coupling is not advisable if there are internal Tyr residues in the peptide. EDCI coupling is not advisable when internal Glu, Asp, or Lys residues are present. Glutaraldehyde coupling may not be appropriate if there are internal Lys residues in the peptide. Table 9.3.3 lists the principal reagents used for peptide-protein conjugation and the functional groups involved.

For example, if the carboxy-terminal sequence of a protein is

1 2 3 4 5 6 7 8 9 10  
S Y G R N Q A E C Q—COOH

then coupling via MBS by adding a Cys residue to the N terminus may not be appropriate because of the Cys at position 9 (see Table 9.3.1 for single-letter codes). In this case, it may be preferable to couple the peptide via the N terminus using glutaraldehyde. However, if the Cys at position 9 is known to be part of a disulfide loop in the native protein, it may be better to couple with MBS through the naturally occurring Cys at position 9. Protocols for the coupling methods discussed here are presented in UNIT 9.4.

#### LITERATURE CITED

- Chou, P.Y. and Fasman, G.D. 1974. Prediction of protein conformation. *Biochemistry* 13:222-245.
- Corrigan, A.J. and Huang, P.C. 1982. A basic micro-computer program for plotting the secondary structure of proteins. *Comput. Programs Biomed.* 3:163-168.
- Dyson, H.J., Cross, K.J., Houghten, R.A., Wilson, I.A., Wright, P.E., and Lerner, R.A. 1985. The immunodominant site of a synthetic immunogen has a conformational preference in water for a type-II reverse turn. *Nature (London)* 318:480-483.

- Hopp, T.P. and Woods, K.R. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. U.S.A.* 78:3824-3828.
- Hopp, T.P. and Woods, W.R. 1983. A computer program for predicting protein antigenic determinants. *Mol. Immunol.* 20:483-489.
- Karplus, P.A. and Schulz, G.E. 1985. Prediction of chain flexibility in proteins: A tool for the selection of peptide antigens. *Naturwissenschaften* 72:212-213.
- Kyte, J. and Doolittle, R.F. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132.
- Van Regenmortel, M.H.V., Briand, J.P., Muller, S., and Plaué, S. 1988. Synthetic polypeptides as antigens. In *Laboratory Techniques in Biochemistry and Molecular Biology*, Vol. 19. (R.H. Burdon and P.H. von Knippenberg, eds.). Elsevier, Amsterdam and New York.
- von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucl. Acids Res.* 14:4683-4690.
- Westhof, E., Altschuh, D., Moras, D., Bloomer, A.C., Mondragon, A., Klug, A., and Van Regenmortel, M.H.V. 1984. Correlation between segmental mobility and the location of antigenic determinants in proteins. *Nature (London)* 311:123-126.

#### KEY REFERENCE

Van Regenmortel et al., 1988. See above.

*Presents detailed discussion of selection of peptides for preparing anti-peptide sera reactive with native proteins.*

Contributed by W. Lee Maloy  
Magainin Sciences  
Plymouth Meeting, Pennsylvania

John E. Coligan  
National Institute of Allergy and  
Infectious Diseases  
Bethesda, Maryland